Noise Reduction Using Frequency-Warped FIR Wiener Filter

Aida Shamsa, Seyed Ghorshi, Marjan Joorabchi
School of Science and Engineering, Sharif University of Technology- International Campus
Kish Island, Iran
{shamsa, joorabchi}@kish.sharif.edu, ghorshi@sharif.edu

***Abstract* -Speech signals can be degraded by different types of additive noises which degrade their quality and intelligibility. Therefore, many noise reduction processes have been proposed to remove the additive noise from a speech signal. Noise reduction methods can be used to substantially reduce noise and interference and, hence improve the perceived quality and intelligibility of a speech signal. Wiener filter is a classical noise reduction method that is widely used for removing noises from the speech signal. In this paper, Wiener filter method based on frequency warping has been proposed for noise reduction process. The applied frequency warped method to the FIR filter is useful in noise reduction of speech signals and yields better speech perceived quality. Furthermore, the results of All-pole FIR Warped Wiener filter are compared to FIR Wiener filter. In addition, the Perceptual Evaluation of Speech Quality (PESQ) of objective quality score for each method is compared.**

Keywords—Noise Reduction; Wiener filter; Warped FIR filter; All-pole filter; PESQ

I. Introduction

Speech signals play an important role to convey messages. Due to this reason, the information of the speech should be preserved. The quality of speech signal can be degraded by interfering noise in a communication channel. When the quality and intelligibility of speech signal are degraded, it will be difficult to understand and identify the speech signal. Consequently, it is aimed to reduce the noise in a way that the speech can be able to convey its carrying information. In general, speech enhancement in noisy environments improves the quality and intelligibility of speech and reduces communication fatigue. Noise reduction can be used in wide range of applications such as hands-free and mobile phones, teleconferencing, in-car cabin communication, etc.

Noise reduction could be divided into two categories. The first category is spectral restoration based on estimation theory and the second one is statistical model-based method. Spectral restoration methods consider the noise reduction as a spectral estimation problem. It means that these methods intend to estimate the spectrum of clean speech out of its noisy version [1]. Ephraim and Malah proposed optimal spectral amplitude estimators by utilizing the statistical estimation theory [2]. Their paper has been cited as one of the primary attempts for employing statistical theory for noise reduction purposes. Statistical Model-based method for noise reduction also considers the procedure of noise reduction as an estimation problem. However, in this category, a mathematical model is utilized to model the speech production process. Moreover, the parameters of speech are estimated in this model have fewer dimensions than signal space itself [3].

Wiener filters are used for enhancement of speech signal which is contaminated with background noise. The procedure of speech enhancement for noise reduction aims to minimize the power of additive noise by Wiener filtering [4]. Consider a noisy speech signal $y(m)$ that is modelled as

$$y(m) = x(m) + n(m) \tag{1}$$

where $x(m)$ is the clean speech signal and $n(m)$ is the additive background noise. The additive noise can exist in the form of environmental noise such as Babble (BAB), Highway (HWY), and Large Crowd (LCR) or in the form of white Gaussian noise. These types of noises can obviously make difficulties in the receiver side of the speech transmission. In this paper White Gaussian Noise (WGN) has been used as the common noise.

$$x_{WGN}[n] \sim N(0, \sigma^2) \text{ for } -\infty < n < \infty \tag{2}$$

As it is known, the FIR model of Wiener filter classification is one of the speech enhancement approaches which is used for noise reduction. In the FIR model of Wiener filter classification, the output is given as

$$\hat{x}(m) = \sum_{k=0}^{p-1} w_k y(m-k) \tag{3}$$

where $\hat{x}(m)$ is the estimation of the desired signal $x(m)$ by the least square error estimation procedure, $\{w_k\}$ are the filter coefficients and $y(m)$ is the noisy observed speech signal.

In this paper the aim is to implement the FIR warped Wiener filter with an all-pole system, while in the previous works, the warping technique was implemented with all-pass system for de-noising a noisy speech signal. The FIR warping is one of the beneficial works for the noise reduction and thus the FIR warped Wiener filter is applied to improve the speech quality. According to binaural noise reduction technique that uses frequency-warped FIR filters, the obtained spectrum can be similar to the auditory representation [5]. The purpose of this study not only

978-1-5090-0206-1/15 $31.00 © 2015 IEEE
DOI 10.1109/EMS.2015.26

preserves the particularity of previous methods but also raises the noise reduction results and the Perceptual Evaluation of Speech Quality (PESQ) scores.

It was found that the human ear perceives frequencies in non-linear form. For this reason, the non-uniform frequency techniques such as frequency warping were taken into account in order to achieve the frequency transformation in resemblance to the frequency perception in human auditory system. Approximately, the frequency warping yields robust solution in numerical difficult problems of filter designing [6]. Warped Digital Signal Processing (DSP) system can be accurately designed based on frequency resolution of human hearing which is observed in quality gain [7, 8, 9].

The rest of this paper is organized as follows. In section II the all-pole FIR Warped filter is described. Section III presents the all-pole FIR warped Wiener filter. Experimental results are presented in section IV and finally section V concludes the paper.

II. All-Pole FIR Warped Filter

The FIR filter's transfer function [4, 10] is expressed as follows

$$\mathrm{H(z)} = \sum_{\mathrm{k=0}}^{\mathrm{M}} \mathrm{b_k z^{-k}} \tag{4}$$

where M is filter's order and $\{b_k\}$ are the FIR filter coefficients. An all-pole filter, $A(z)$, is used as a transfer function in the frequency warping procedure and is given by (5)

$$A(z) = \frac{1}{1+\sum_{k=1}^{p} a_k z^{-k}} \tag{5}$$

where $\{a_k\}$ are the coefficients of the all-pole system with the order of P. The FIR filter's coefficients have been calculated by Linear Prediction Coefficient (LPC) method [11] because it is considered that the all-pole coefficients are as the same as FIR coefficients. Also, the all-pole filter has the same order as the FIR filter.

In order to obtain the transfer function of warped FIR, the transfer function of all-pole filter, $A(z)$, in (5) must be replaced with z^{-1} factor in FIR transfer function in (4). By doing this procedure, the transfer function of Warped FIR (WFIR) can be expressed as (7),

$$z^{-1} \rightarrow \frac{1}{1+\sum_{k=1}^{p} a_k z^{-k}} \tag{6}$$

$$H_{WFIR}(z) = \sum_{K=0}^{M} \beta_k \{A(z)\}^k = \sum_{K=0}^{M} \beta_k \left\{\frac{1}{1+\sum_{k=1}^{p} a_k z^{-k}}\right\}^k \tag{7}$$

where $\{\beta_k\}$ are the warp coefficients [7] and M is WFIR filter's order.

III. All-Pole Fir Warped Wiener Filter

Since the Wiener filter method is used for noise reduction, the described frequency warping was exerted to the FIR Wiener filter.

As it is known, Wiener filter is commonly used for estimating the output signal based on Minimum Mean Square Error (MMSE). The output of Wiener filter equation in frequency domain is given in (8)

$$\hat{X}(f) = W(f)Y(f) \tag{8}$$

where $\hat{X}(f)$ is the Wiener filter output, $Y(f)$ is the input signal and $W(f)$ is the filter frequency response. Considering this fact, the noisy signal in (1) in frequency domain can be expressed as

$$Y(f) = X(f) + N(f) \tag{9}$$

where $Y(f)$, $X(f)$ and $N(f)$ are the noisy signal, clean signal and noise spectra respectively.

The estimated speech signal can be obtained by solving the Minimum Mean Square Error (MMSE) equation that is employed to calculate the coefficients of Wiener filter or the frequency response of Wiener filter. The power spectrum of speech signal that is utilized in calculation of Wiener coefficients or frequency response of Wiener filter can be achieved by taking the Fourier Transform of the autocorrelation equation as follows,

$$P_{xx}(f) = E[X(f)X^{*}(f)] = \sum_{k=-\infty}^{\infty} r_{xx}(k)e^{-j2\pi k} \tag{10}$$

where $P_{xx}(f)$ is the power spectrum of the speech signal. Hence, for a noisy speech signal, the Wiener filter frequency response is expressed as (11)

$$W(f) = \frac{P_{xx}(f)}{P_{xx}(f)+P_{nn}(f)} \tag{11}$$

where $P_{xx}(f)$ and $P_{nn}(f)$ are the power spectrums of the speech signal and noise respectively. Also, considering the time domain of output signal of Wiener filter in (3), the Wiener filter coefficients $\{w_k\}$ and the noisy speech signal $y(m)$ have been used for estimating the filter output $\hat{x}(m)$. Therefore, in order to utilize the warping method in the time domain Wiener filter function, all the z^{-1} factors will be replaced with (6). This warping process helps to reduce the additive noise from the speech signal. In addition, the compatibility of frequency warping with the human perception system can ease the realization of the noisy speech after filtering the signal compared with traditional Wiener filter method.

IV. The Experimental Results

The proposed method for noise reduction has been tested on 100 sentences, which are contaminated artificially by additive White Gaussian Noise.

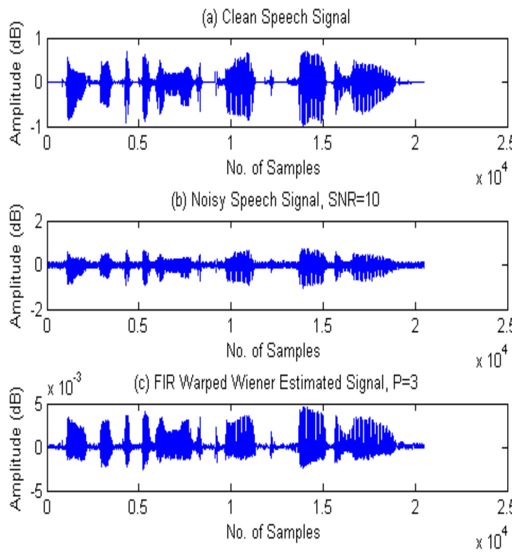

Figure 1. (a) Clean speech signal, (b) noisy speech signal with $SNR = 10\ dB$, (c) estimated speech signal using FIR warped Wiener filter with $P = 3$.

These sentences are spoken by female speakers and selected randomly from TIMIT database [12]. The sampling frequency f_s, of speech signal is 16 kHz and since the speech signals are non-stationary it is desired to employ in framing to the speech signal in order to achieve stationary processes after reading the original speech signal. It is also tried to compare all-pole FIR warped Wiener filter with FIR Wiener filter method in terms of spectrograms of de-noised speech signal to observe the efficiency of these approaches for noise reduction.

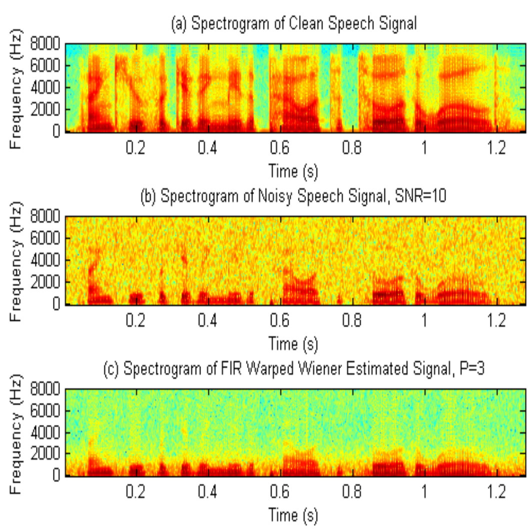

Figure 2. (a) Spectrograms of clean speech signal, (b) noisy speech signal with $SNR = 10\ dB$, (c) estimated speech signal using FIR warped Wiener filter with $P = 3$.

The estimated speech signal with FIR warped Wiener filter by an all-pole system with order $P = 3$ from the noisy speech signal with $SNR = 10\ dB$ is illustrated in Fig. 1. In addition, the spectrograms of estimated speech signal using FIR warped Wiener filter by an all-pole system with order $P = 3$ from the noisy speech signal with $SNR = 10\ dB$ is illustrated in Fig. 2. Fig. 3 also shows the spectrograms of clean speech, the noisy speech and the estimated Wiener filter speech signals with $SNR = 10\ dB$ using the same filter order i.e. $P = 3$.

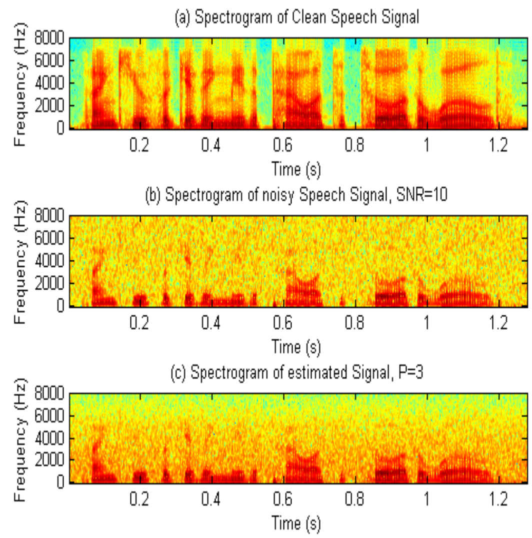

Figure 3. (a) Spectrograms of clean speech signal, (b) noisy speech signal with $SNR = 10\ dB$, (c) estimated speech signal using Wiener filter with $P = 3$.

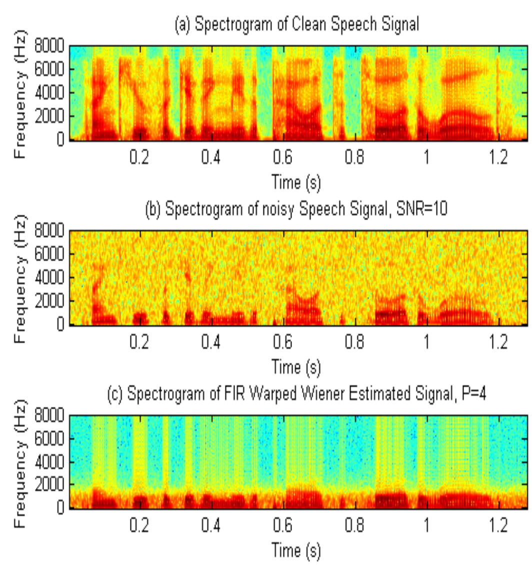

Figure 4. (a) Spectrograms of clean speech signal, (b) noisy speech signal with $SNR = 10\ dB$, (c) estimated speech signal using FIR warped Wiener filter with $P = 4$.

By comparing the estimated speech signal using an all-pole FIR warped Wiener filter (Fig. 2) with the estimated speech signal using Wiener filter (Fig. 3) with the same SNR and filter order, it can be found that the proposed method (FIR warped Wiener filter) outperforms the Wiener filter for noise reduction. Moreover, in order to obtain better results and observe the efficiency of these approaches for noise reduction we have increased the filter order P for both all-pole FIR warped Wiener filter and Wiener filter.

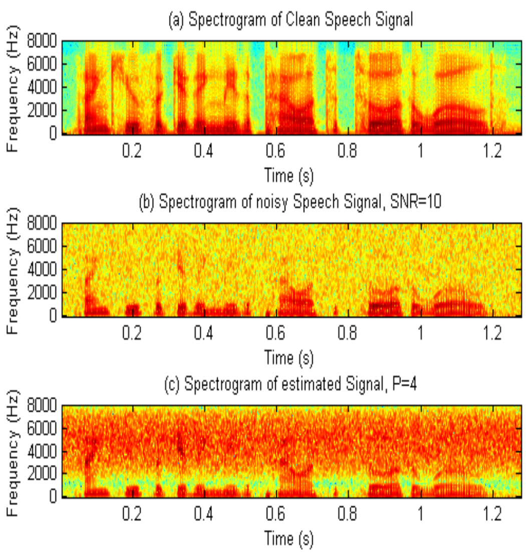

Figure 5. (a) Spectrograms of clean speech signal, (b) noisy speech signal with $SNR = 10\ dB$, (c) estimated speech signal using Wiener filter with $P = 4$.

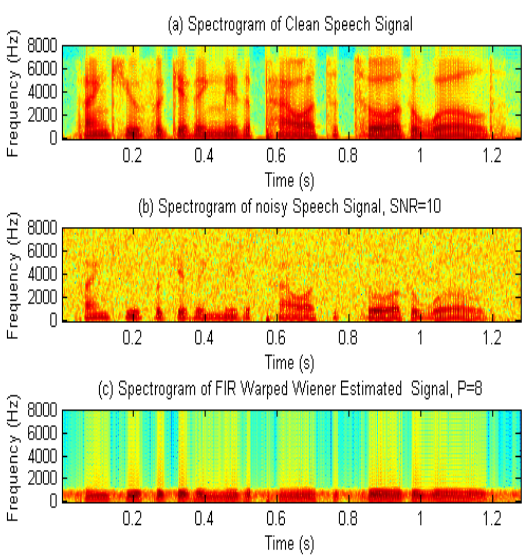

Figure 6. (a) Spectrograms of clean speech signal, (b) noisy speech signal with $SNR = 10\ dB$, (c) estimated speech signal using FIR warped Wiener filter with $P = 8$.

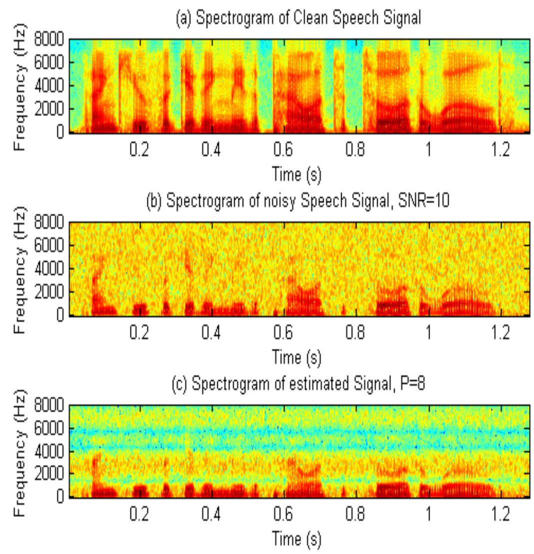

Figure 7. (a) Spectrograms of clean speech signal, (b) noisy speech signal with $SNR = 10\ dB$, (c) estimated speech signal using Wiener filter with $P = 8$.

It should be noted that changing the order of FIR filter also changes the order of all-pole filter. Figs. 4 and 5 show the spectrograms of estimated speech signal using FIR warped Wiener filter by an all-pole system with order $P = 4$ from the noisy speech signal with $SNR = 10\ dB$ and the spectrograms of clean speech, the noisy speech and the estimated Wiener filter speech signals with $SNR = 10\ dB$ using the same filter order respectively. The last examining between the filters in terms of comparing the spectrograms of clean, noisy and estimated speech signals are exhibited in Figs. 6 and 7, where the order of the filters has been increased to $P = 8$. By increasing the order of the filter the quality and intelligibility of speech signal have been improved and also the effect of noise on speech signal has been reduced. As complementary evaluation criteria, the comparison between objective quality score from the Perceptual Evaluation of Speech Quality (PESQ) for each method has been provided in Table I. The results of Table I show that the Mean Opinion Score (MOS) of PESQ of FIR warped Wiener Filter provides better result compared to Wiener Filter.

TABLE I. CAMPARISON RESULT OF OBJECTIVE QUALITY SCORE FROM PESQ BETWEEN FIR WARPED WIENER FILTER AND WIENER FILTER WITH SNR=10 dB.

Filter Order $P = 3$		Filter Order $P = 4$		Filter Order $P = 8$	
FIR warped Wiener Filter	Wiener Filter	FIR warped Wiener Filter	Wiener Filter	FIR warped Wiener Filter	Wiener Filter
2.16	1.73	2.63	1.86	2.68	2.15

It can also be seen from Table I that as the order of filter increases the MOS of PESQ of both, FIR warped Wiener filter and Wiener filter increases too. Also, it can be observed that for $P = 8$ a significant improvement is achieved in Wiener filter compared with other filter orders. Therefore, it is concluded that the performance of FIR warped Wiener filter to reduce the effect of noise on the speech signal is quite better compared with Wiener filter.

V. Conclusion

This paper discussed the FIR warped Wiener filter for reducing the noise from speech signals. Frequency warping was the major idea of this paper. This method was exerted in Wiener filter. First, the FIR filter was warped with an all-pole filter and then the warped FIR filter includes in Wiener filter in order to remove the noise from the speech signal and tailored in the way that human perception system receives signals. Ultimately, the estimated signal of the proposed method under the same value of SNR not the same filter order was obtained. This method has been done for both, Wiener filter and FIR warped Wiener filter and the proposed technique demonstrated better results of noise reduction in comparison with the traditional Wiener filter.

References

[1] S. F. Boll, "Suppression of acoustic noise in speech using Spectral Subtraction," IEEE Trans. Acoust., Speech, Signal Proccessing, vol. 27, pp. 113-120, April 1979.

[2] Y. Ephraim, and D. Malah, "Speech enhancement using a minimum mean square error log-spectral amplitude estimator," IEEE Trans. Acoust., Speech, Signal Processing, vol. 33, pp. 443-445, April 1985.

[3] Y. Ephraim, "Statistical-Model-Based Speech Enhancement Systems," Proc. IEEE, vol. 80, pp. 1526-1555, October 1992.

[4] S. V. Vaseghi, Multimedia Signal Processing: Theory and Applications in Speech, Music and Communications, England: John Wiley & Sons Ltd, 2007.

[5] J. I. Marin-Hurtado, and D. V. Anderson, "Binaural noise reduction using frequency-warped FIR filters," INTERSPEECH, 2012.

[6] C. Asavathiratham, P. E. Beckmann, and A. V. Oppenheim, "Frequency warping in the design and implementation of fixed-point audio equalizers," Proc. IEEE Workshop Appl. Signal Process. Audio Acoust. (WASPAA 99), IEEE Press, pp. 55-58, October 1999.

[7] A. Härmä, M. Karjalainen, L. Savioja, V. Välimäki, U. K. Laine, and J. Huopaniemi, "Frequency-warped signal processing for audio applications," Journal of Audio Eng. Soc. (JAES), vol. 48, Issue 11, pp. 1011-1031, November 2000.

[8] A. Härmä, "Implementation of frequency-warped recursive filters," Signal Processing, vol. 80, pp. 543-548, March 2000.

[9] A. Farina, A. Bellini, and E. Armelloni, "Implementation of cross-talk canceling filters with warped structures -subjective evaluation of the loudspeaker reproduction of stereo recordings," Proc. SHARC2000, 2000.

[10] X. Huang, A. Acero, and H. W. Hon, Spoken Language Processing: A Gide to Theory, Algorithm, and System Development, Prentice Hall PTR, 2001.

[11] L. R. Rabiner, and R. W. Schafer, "Introduction to Digital Speech Processing," Foundations and Trends in Signal Processing, vol. 1, pp. 1-194, 2007.

[12] TIMIT Dictionary avaiable: http://www.ldc.upenn.edu/.